AD-A118 974 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH F/G 12/1
POLYNOMIAL APPROXIMATION: THE WEIERSTRASS APPROXIMATION THEORFM--ETC(U)
AUG 82 S J NICHOLS
UNCLASSIFIED AFIT/CI/NR/82-46T NL

1 of 1
AD A118974

END
DATE
FILMED
10 82
DTIC

UNCLASS
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFIT/CI/NR/82-46T	2. GOVT ACCESSION NO. AD-A118 974	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Polynomial Approximation: The Weierstrass Approximation Theorem		5. TYPE OF REPORT & PERIOD COVERED THESIS/~~DISSERTATION~~
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Shirley Jo Nichols		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS AFIT STUDENT AT: University of Hawaii		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS AFIT/NR WPAFB OH 45433		12. REPORT DATE August 1982
		13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASS
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

APPROVED FOR PUBLIC RELEASE: IAW AFR 190-17

Lynn Wolaver
LYNN E. WOLAVER
Dean for Research and
Professional Development
AFIT, Wright-Patterson AFB OH

30 AUG 1982

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

ATTACHED

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASS
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

82 09 07 411

ABSTRACT

In this paper we will look at three proofs of the Weierstrass Approximation Theorem. The first proof is in much the same form in which Weierstrass originally proved his theorem. The next is due to Lebesgue. It is by far the easiest proof to follow, with only a minimum knowledge of analysis required. The last arises from probability and uses the Bernstein polynomials.

Secondly we look at a generalization of this theorem, called the Stone-Weierstrass Theorem. This generalization was inspired by modern developments in mathematics. The theorem deals with functions on a general compact space rather than on a closed interval.

A

AFIT/CI/NR/82-46T

AFIT RESEARCH ASSESSMENT

The purpose of this questionnaire is to ascertain the value and/or contribution of research accomplished by students or faculty of the Air Force Institute of Technology (ATC). It would be greatly appreciated if you would complete the following questionnaire and return it to:

AFIT/NR
Wright-Patterson AFB OH 45433

RESEARCH TITLE: Polynomial Approximation: The Weierstrass Approximation Theorem

AUTHOR: Shirley Jo Nichols

RESEARCH ASSESSMENT QUESTIONS:

1. Did this research contribute to a current Air Force project?

( ) a. YES ( ) b. NO

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFIT had not?

( ) a. YES ( ) b. NO

3. The benefits of AFIT research can often be expressed by the equivalent value that your agency achieved/received by virtue of AFIT performing the research. Can you estimate what this research would have cost if it had been accomplished under contract or if it had been done in-house in terms of manpower and/or dollars?

( ) a. MAN-YEARS ________ ( ) b. $________

4. Often it is not possible to attach equivalent dollar values to research, although the results of the research may, in fact, be important. Whether or not you were able to establish an equivalent value for this research (3. above), what is your estimate of its significance?

( ) a. HIGHLY SIGNIFICANT ( ) b. SIGNIFICANT ( ) c. SLIGHTLY SIGNIFICANT ( ) d. OF NO SIGNIFICANCE

5. AFIT welcomes any further comments you may have on the above questions, or any additional details concerning the current application, future potential, or other value of this research. Please use the bottom part of this questionnaire for your statement(s).

NAME GRADE POSITION

ORGANIZATION LOCATION

STATEMENT(s):

FOLD DOWN ON OUTSIDE - SEAL WITH TAPE

AFIT/NR
WRIGHT-PATTERSON AFB OH 45433

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE. $300

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST CLASS PERMIT NO. 73236 WASHINGTON D.C.

POSTAGE WILL BE PAID BY ADDRESSEE

AFIT/ DAA
Wright-Patterson AFB OH 45433

FOLD IN

82-46T

POLYNOMIAL APPROXIMATION:

THE WEIERSTRASS APPROXIMATION THEOREM

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

IN MATHEMATICS

AUGUST 1982

By

Shirley Jo Nichols

Thesis Committee:

Lawrence J. Wallen, Chairman
Scott W. Brown
Thomas B. Hoover

We certify that we have read this thesis and that in our opinion it is satisfactory in scope and quality as a thesis for the degree of Master of Arts, in Mathematics.

THESIS COMMITTEE

Lawrence J Wallen
Chairman

Thomas B. Hoover

Scott W. Brown

TABLE OF CONTENTS

INTRODUCTION 1

I. PRELIMINARY REMARKS. 2

II. PROOF DUE TO WEIERSTRASS. 3

III. PROOF DUE TO LEBESGUE. 8

IV. PROOF DUE TO BERNSTEIN. 13

V. THE STONE-WEIERSTRASS THEOREM. 19

BIBLIOGRAPHY 25

INTRODUCTION

In 1885 Karl Weierstrass showed that every continuous function in a closed interval can be uniformly approximated by polynomials. This result is called the Weierstrass Approximation Theorem.

Many mathematicians have since found new proofs of this theorem, using techniques which arise from their particular fields of interest. Some proofs give extra information about the approximating polynomial, others do not. Some require advanced knowledge of various fields, others can be understood with a basic knowlege of analysis.

In this paper we will look at three proofs of the Weierstrass Approximation Theorem. The first proof is in much the same form in which Weierstrass originally proved his theorem. The next is due to Lebesgue. It is by far the easiest proof to follow, with only a minimum knowledge of analysis required. The last arises from probability and uses the Bernstein polynomials.

Secondly we look at a generalization of this theorem, called the Stone-Weierstrass Theorem. This generalization was inspired by modern developments in mathematics. The theorem deals with functions on a general compact space rather than on a closed interval.

I. PRELIMINARY REMARKS

Before we actually begin looking at the Weierstrass Approximation Theorem we consider a preliminary lemma. It shows that it is sufficient to prove the Weierstrass Approximation Theorem on the closed interval $[0,1]$.

Lemma

It is sufficient to prove the Weierstrass Approximation Theorem for the special case $[a,b] = [0,1]$.

Proof

If $a \leq x \leq b$, and $x = a + (b - a)y$, then $0 \leq y \leq 1$. Let f be a continuous function on $[a,b]$. Set $g(y) = f(x)$. Then g is a continuous function on $[0,1]$. By hypothesis, there exists a polynomial p such that $|g(y) - p(y)| < \varepsilon$ for $0 \leq y \leq 1$. Then

$$\left|f(x) - p\left(\frac{x - a}{b - a}\right)\right| < \varepsilon$$

for $a \leq x \leq b$, and $p\left(\frac{x - a}{b - a}\right)$ is a polynomial.

II. PROOF DUE TO WEIERSTRASS

In 1885, Weierstrass proposed and solved the following: If $f(x)$ is a continuous function, is it possible to make the error of approximation arbitrarily small by increasing the degree of the approximating polynomial? By looking at the three theorems in this section we see that the answer to this question is yes; and we will show this in much the same form as Weierstrass first proved it. These theorems require knowledge of some graduate-level analysis.

Remark

It is sufficient to prove the Weierstrass Approximation Theorem for the special case $f(1) = f(0) = 0$. For if we were to prove the theorem using these conditions, we see that for any continuous f on $[0,1]$ we could let $g(x) = f(x) - f(0) - x[f(1) - f(0)]$ for $0 \leq x \leq 1$. Then $g(0) = g(1) = 0$, and if g can be obtained as the limit of a uniformly convergent sequence of polynomials, it is clear that the same is true for f, since $f - g$ is a polynomial.

Theorem 1

If f is a continuous function on $[0,1]$ and if g_n is a family of functions such that:

a) $g_n \geq 0$;

b) For each $\varepsilon > 0$ there exists a $\delta > 0$ such that $\int_{|t|>\delta} g_n(t)dt < \varepsilon$ if n is large;

c) $\int_{-\infty}^{\infty} g_n(t)dt = 1;$

d) g_n is even,

then $(g_n * f)(x) \to f(x)$ uniformly on $[0,1]$, where $(F * G)(x)$ is defined by $(F * G)(x) = \int_{-\infty}^{\infty} F(x - t)G(t)dt$, if this integral exists.

<u>Proof</u>

Assume by the previous remark that $f(0) = f(1) = 0$.

Since f is continuous on $[0,1]$, a closed interval, f is uniformly continuous on $[0,1]$. Furthermore, we define $f(x)$ to be zero for x outside $[0,1]$. Then f is uniformly continuous on the whole real line.

Given $\varepsilon > 0$, we choose $\delta_1 > 0$ such that $|y - x| < \delta_1$ implies that $|f(y) - f(x)| < \frac{\varepsilon}{2}$, which is permissible by uniform continuity. Choose $\delta_o > 0$ so that $\int_{|t|>\delta_o} g_n(t)dt < \frac{\varepsilon}{4M}$, where $M = \sup|f(x)|$, for $n > N_o$. Let $\delta = \min(\delta_o,\delta_1)$. First we note that

$$\begin{aligned}(g_n * f)(x) &= \int_{-\infty}^{\infty} g_n(x - u)f(u)du \\ &= \int_{-\infty}^{\infty} g_n(u - x)f(u)du \quad \text{(using the evenness of } g).\end{aligned}$$

Let $u - x = t$. Then $u = x + t$, so that

$$(g_n * f)(x) = \int_{-\infty}^{\infty} g_n(t)f(x + t)\, dt.$$

We see that for $0 \leq x \leq 1$

$$
\begin{aligned}
|(g_n * f)(x) - f(x)| &= \left|\int_{-\infty}^{\infty} g_n(t) f(x+t)\,dt - f(x)\right| \\
&= \left|\int_{-\infty}^{\infty} f(x+t) g_n(t)\,dt - f(x) \int_{-\infty}^{\infty} g_n(t)\,dt\right| \\
&= \left|\int_{-\infty}^{\infty} [f(x+t) - f(x)] g_n(t)\,dt\right| \\
&\le \int_{-\infty}^{\infty} |f(x+t) - f(x)| g_n(t)\,dt \\
&= \int_{-\delta}^{\delta} |f(x+t) - f(x)| g_n(t)\,dt + \int_{|t|>\delta} |f(x+t) - f(x)| g_n(t)\,dt \\
&< \int_{-\delta}^{\delta} \frac{\varepsilon}{2} g_n(t)\,dt + \int_{|t|>\delta} 2M g_n(t)\,dt \\
&= \frac{\varepsilon}{2} \int_{-\delta}^{\delta} g_n(t)\,dt + 2M \int_{|t|>\delta} g_n(t)\,dt \\
&\le \frac{\varepsilon}{2} \cdot 1 + 2M \cdot \frac{\varepsilon}{4M} = \varepsilon, \quad \text{for all large enough } n.
\end{aligned}
$$

Theorem 2

If $Q_n(x) = c_n(1 - x^2)^n$ for $-1 \le x \le 1$, and $Q_n(x) = 0$ otherwise, for $n = 1, 2, 3, \ldots,$ then the c_n's can be chosen so that the following conditions hold:

a) $Q_n(x)$ is even;

b) $\int_{-\infty}^{\infty} Q_n(t)\,dt = 1;$

c) $Q_n(x) \ge 0;$

d) For each $\varepsilon > 0$ there exists a $\delta > 0$ such that $\int_{|t|>\delta} Q_n(t)\,dt < \varepsilon$ if n is large.

Proof

To see that $Q_n(x)$ is even, substitute $-x$ for x. Then $Q_n(-x) = c_n(1 - (-x)^2)^n = c_n(1 - x^2)^n = Q_n(x)$.

$\int_{-\infty}^{\infty} (1 - t^2)^n dt$ is obviously finite, so we can let

$c_n = \dfrac{1}{\int_{-\infty}^{\infty} (1 - t^2)^n dt}$. Then $\int_{-\infty}^{\infty} Q_n(t)dt = 1$.

$(1 - x^2)^n \geq 0$ for $-1 \leq x \leq 1$. This implies that $c_n \geq 1$ for all n. Thus $Q_n(x) \geq 0$.

We need to find an upper bound on c_n. Since $|x| > x^2$ for $0 \leq |x| \leq 1$, we have that $(1 - x^2)^n > (1 - |x|)^n$.

Then

$$\begin{aligned}\int_{-\infty}^{\infty} (1 - x^2)^n dx &= \int_{-1}^{1} (1 - x^2)^n dx > 2\int_{0}^{1} (1 - |x|)^n dx \\ &= \frac{-2(1 - x)^{n+1}}{n + 1}\Big]_0^1 = \frac{-2(1 - 1)^{n+1}}{n + 1} - \frac{-2(1 - 0)^{n+1}}{n + 1} \\ &= \frac{2}{n + 1}.\end{aligned}$$

Since

$$\begin{aligned}1 = \int_{-\infty}^{\infty} Q_n(x)dx &= \int_{-1}^{1} c_n(1 - x^2)^n dx \\ &= c_n\int_{-1}^{1} (1 - x^2)^n dx > \frac{2c_n}{n + 1},\end{aligned}$$

we have that $c_n < \frac{n + 1}{2}$.

For any $\delta > 0$, if $\delta \leq |x| \leq 1$ then $\delta^2 \leq x^2$. Thus $1 - \delta^2 \geq 1 - x^2$ and $(1 - \delta^2)^n \geq (1 - x^2)^n$ for $n = 1, 2, 3, \ldots$. If $Q_n(x) = c_n(1 - x^2)^n$ where $c_n < \frac{n + 1}{2}$, then

$Q_n(x) = c_n(1 - x^2)^n < \frac{n + 1}{2}(1 - x^2)^n \leq \frac{n + 1}{2}(1 - \delta^2)^n$ for $\delta \leq |x| \leq 1$.

Given $\varepsilon > 0$, we have

$$\int_{|t|>\delta} Q_n(t)dt < \int_{|t|>\delta} \frac{n+1}{2}(1-\delta^2)^n dt$$
$$= \frac{n+1}{2}(1-\delta^2)^n \int_{|t|>\delta} dt$$
$$= \frac{n+1}{2}(1-\delta^2)^n \, 2\int_{\delta}^{1} dt$$
$$= (n+1)(1-\delta^2)^n(1-\delta).$$

This last quantity goes to zero, as $n \to \infty$. Thus $\int_{|t|>\delta} Q_n(t)dt < \varepsilon$ if n is large.

<u>*Theorem 3*</u>

Let f be as in Theorem 1. Then $P_n(x) = (Q_n * f)(x)$ is a polynomial and thus $P_n(x)$ is a family of polynomials which approximates f uniformly.

<u>Proof</u>

Our assumptions about f show, by a simple change of variables, that

$$P_n(x) = (Q_n * f)(x) = \int_{-\infty}^{\infty} Q_n(u)f(x+u)du$$
$$= \int_{-1}^{1} Q_n(u)f(x+u)du = \int_{-x}^{1-x} f(x+u)Q_n(u)du$$
$$= \int_{0}^{1} f(t)Q_n(t-x)dt,$$

where $x + u = t$, or $u = t - x$. Thus $P_n(x) = \int_0^1 f(t)Q_n(t-x)dt$, which is clearly a polynomial in x. The final conclusion follows from Theorems 1 and 2.

III. PROOF DUE TO LEBESGUE

There is a multitude of proofs of the Weierstrass Approximation Theorem. However, one of the simplest and most direct proofs is due to Lebesgue, in 1898. (Because of its simplicity, this proof is suitable for undergraduate mathematics majors.) The basic method of the proof is first to establish that the function $f(x)$ may be approximated arbitrarily closely by a broken line. That this may be done follows from the fact that $f(x)$ is continuous on a closed interval, and hence uniformly continuous. The second step is to show that the broken line may be approximated arbitrarily closely by a polynomial.

Our proof will be presented in the form of three lemmas.

<u>Lemma 1</u>

Let f be a continuous function on $[0,1]$. Then f can be uniformly approximated by piecewise linear functions.

<u>Proof</u>

Pick δ such that $|f(u) - f(v)| < \frac{\varepsilon}{2}$ if $|u - v| < \delta$ (by uniform continuity). Pick x_i, $i = 0, 1, \ldots, k$ with $0 = x_0 < x_1 < \ldots < x_k = 1$ so that $|x_{j+1} - x_j| < \delta$. Define φ such that $\varphi(x_j) = f(x_j)$ for $j = 0, 1, \ldots, k$ and such that φ is linear between the x_j's. Then if $x_j < x < x_{j+1}$,

$$\varphi(x) = \frac{f(x_{j+1}) - f(x_j)}{x_{j+1} - x_j}(x - x_j) + f(x_j).$$

Then

$$|f(x) - \varphi(x)| = \left|f(x) - \left[\frac{f(x_{j+1}) - f(x_j)}{x_{j+1} - x_j}(x - x_j) + f(x_j)\right]\right|$$

$$= \left|f(x) - \left[\frac{x - x_j}{x_{j+1} - x_j}f(x_{j+1}) + \left\{1 - \left(\frac{x - x_j}{x_{j+1} - x_j}\right)\right\}f(x_j)\right]\right|$$

$$= \left|\frac{x - x_j}{x_{j+1} - x_j}[f(x) - f(x_{j+1})] + \left(1 - \frac{x - x_j}{x_{j+1} - x_j}\right)[f(x) - f(x_j)]\right|$$

$$\leq \left(\frac{x - x_j}{x_{j+1} - x_j}\right)|f(x) - f(x_{j+1})| + \left(1 - \frac{x - x_j}{x_{j+1} - x_j}\right)|f(x) - f(x_j)|$$

$$< 1 \cdot \frac{\varepsilon}{2} + 1 \cdot \frac{\varepsilon}{2} = \varepsilon.$$

<u>Definition</u>

For each real number a, define $\varphi_a(x) = \max(x - a, 0)$.

<u>Lemma 2</u>

The function $\varphi(x)$ of Lemma 1 can be written

$$\varphi(x) = f(0) + \sum_{j=0}^{k-1} a_j \varphi_{x_j}(x)$$

for $0 \leq x \leq 1$.

<u>Proof</u>

The function $\varphi(x)$ is clearly linear for $x \notin \{x_0, x_1, \ldots, x_k\}$. All we need to show is that we can find $a_0, \ldots, a_{k-1}$ so that $\varphi(x_j) = f(x_j)$. This is equivalent to solving the following system:

$$\varphi(x_o) = f(0)$$

$$\varphi(x_1) = f(0) + a_o\varphi_{x_o}(x_1) = f(x_1)$$

$$\varphi(x_2) = f(0) + a_o\varphi_{x_o}(x_2) + a_1\varphi_{x_1}(x_2) = f(x_2)$$

$$\vdots \qquad\qquad \vdots$$

$$\varphi(x_k) = f(0) + a_o\varphi_{x_o}(x_k) + a_1\varphi_{x_1}(x_k) + \dots + a_{k-1}\varphi_{x_{k-1}}(x_k) = f(x_k)$$

This system can be trivially solved recursively, proving the lemma.

Remark

It follows from Lemmas 1 and 2 that if we can approximate the functions $\varphi_a(x)$ arbitrarily closely on $[0,1]$ by polynomials, then we can approximate $f(x)$ arbitrarily closely on $[0,1]$ by polynomials. But $\varphi_a(x) = \frac{1}{2}(|x - a| + (x - a))$, so, in reality, we have to approximate $|x - a|$ on $[0,1]$. For this, it clearly suffices to approximate $|x|$ on an arbitrary interval $[-A,A]$. But if we can approximate $|u|$ on $[-1,1]$, that is, if

$$\Big||u| - p(u)\Big| < \frac{\varepsilon}{A} \quad \text{on} \quad -1 \leq u \leq 1,$$

then, setting $u = \frac{x}{A}$, $|x| < A$ and

$$\left|A\left|\frac{x}{A}\right| - Ap\left(\frac{x}{A}\right)\right| < \varepsilon, \quad \text{or}$$

$$\left||x| - Ap\left(\frac{x}{A}\right)\right| < \varepsilon \quad \text{on} \quad |x| < A.$$

Lemma 3

$|x|$ is approximable arbitrarily closely by polynomials on $|x| \leq 1$.

Proof

Consider the function $g(t) = \sqrt{1 - t}$ on $0 \leq t \leq 1$. Write

$$g(t) = T_n(t) + R_n(t)$$

where $T_n(t)$ is the Taylor polynomial of degree n, and $R_n(t)$ is the remainder. We will show that $R_n(t) \to 0$ uniformly on $0 \leq t < 1$.

Write $R_n(t)$ in the Lagrange form

$$R_n(t) = \frac{1}{n!}\int_0^t (t - u)^n g^{(n+1)}(u)du, \quad \text{where}$$
$$g^{(n+1)}(u) = \left(\tfrac{1}{2}\right)\left(\tfrac{1}{2} - 1\right)\left(\tfrac{1}{2} - 2\right) \ldots \left(\tfrac{1}{2} - n\right)\left(1 - u\right)^{-n-\frac{1}{2}}.$$

Then

$$R_n(t) = \frac{\left(\frac{1}{2}\right)\left(\frac{1}{2} - 1\right)\ldots\left(\frac{1}{2} - n\right)}{n!}\int_0^t (t - u)^n(1 - u)^{-n-\frac{1}{2}}du, \quad \text{and}$$
$$|R_n(t)| < \frac{\left(\frac{1}{2}\right)\left(1 - \frac{1}{2}\right)\left(2 - \frac{1}{2}\right) \ldots \left(n - \frac{1}{2}\right)}{n!}\left|\int_0^1 (1 - u)^n(1 - u)^{-n-\frac{1}{2}}du\right|$$
$$= \left(1 - \frac{1}{2n}\right)\left(1 - \frac{1}{2(n - 1)}\right) \ldots \left(1 - \frac{1}{2}\right)\left(\frac{1}{2}\right)\left|\int_0^1 (1 - u)^{-\frac{1}{2}}du\right|$$
$$= \left(1 - \frac{1}{2n}\right)\left(1 - \frac{1}{2(n - 1)}\right) \ldots \left(1 - \frac{1}{2}\right).$$

We need to show that $p_n = \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{4}\right) \ldots \left(1 - \frac{1}{2n}\right) \to 0$ as $n \to \infty$, or that $\log p_n \to -\infty$ as $n \to \infty$. Now

$\log p_n = \log\left(1 - \frac{1}{2}\right) + \log\left(1 - \frac{1}{4}\right) + \ldots + \log\left(1 - \frac{1}{2n}\right)$, and

$\frac{\log(1 + (-h)) - \log 1}{-h} = \frac{\log(1 - h)}{-h} \to +1$ as $h \to 0$.

So if h is small, $\frac{\log(1 - h)}{-h} > \frac{1}{2}$, so $\log(1 - h) < -\frac{h}{2}$. Then, if n is large,

$$\log\left(1 - \frac{1}{2n}\right) + \log\left(1 - \frac{1}{2(n+1)}\right) + \ldots + \log\left(1 - \frac{1}{2(n+p)}\right)$$
$$< -\frac{1}{2}\left(\frac{1}{2n} + \frac{1}{2(n+1)} + \ldots + \frac{1}{2(n+p)}\right).$$

Since the harmonic series diverges, the right side goes to $-\infty$. Thus $p_n \to 0$, and $R_n(t) \to 0$ uniformly for $0 \leq t < 1$. Thus, if n is large, $|R_n(t)| < \varepsilon$ for $0 \leq t < 1$, and $g(t) = |\sqrt{1 - t} - T_n(t)| < \varepsilon$ for $0 \leq t < 1$. The same inequality continues to hold on $0 \leq t \leq 1$ by continuity.

To finish the proof, note that if $|x| \leq 1$, then $0 \leq 1 - x^2 \leq 1$ and $|x| = \sqrt{1 - (1 - x^2)}$. From the argument above, we can find the polynomial p such that $|\sqrt{1 - t} - p(t)| < \varepsilon$ for $0 \leq t \leq 1$, and so

$$\left||x| - p(1 - x^2)\right| = |\sqrt{1 - (1 - x^2)} - p(1 - x^2)| < \varepsilon$$

for $|x| \leq 1$.

IV. PROOF DUE TO BERNSTEIN

This proof utilizes the Bernstein polynomials which arise from the study of probability. In the process of the proof we obtain a specific polynomial that estimates the given function to the desired degree. For this reason this proof is sometimes of interest to applied mathematicians.

Before proceeding with the proof, we need some preliminary computations.

Preliminary Computations

From the binomial theorem we see that for any $p, q \in R$ we have

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p + q)^n \qquad \text{for } n \in I. \qquad \text{(a)}$$

Differentiating with respect to p we obtain

$$\sum_{k=0}^{n} \binom{n}{k} k p^{k-1} q^{n-k} = n(p + q)^{n-1},$$

which implies

$$\sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} p^k q^{n-k} = p(p + q)^{n-1} \quad \text{for } n \in I. \qquad \text{(b)}$$

Differentiating once more we have

$$\sum_{k=0}^{n} \left(\frac{k^2}{n}\right) \binom{n}{k} p^{k-1} q^{n-k} = p(n - 1)(p + q)^{n-2} + (p + q)^{n-1},$$

and so

$$\sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 \binom{n}{k} p^k q^{n-k} = p^2\left(1 - \frac{1}{n}\right)(p + q)^{n-2} + \frac{p}{n}(p + q)^{n-1}. \quad \text{(c)}$$

Now, if $x \in [0,1]$, set $p = x$ and $q = 1 - x$. Then (a), (b) and (c) yield:

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1 - x)^{n-k} = 1,$$

$$\sum_{k=0}^{n} \frac{k}{n}\binom{n}{k} x^k (1 - x)^{n-k} = x,$$

$$\sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 \binom{n}{k} x^k (1 - x)^{n-k} = x^2\left(1 - \frac{1}{n}\right) + \frac{x}{n}.$$

Rewriting these three equations we have:

$$\sum_{k=0}^{n} x^2\binom{n}{k} x^k (1 - x)^{n-k} = x^2,$$

$$-\sum_{k=0}^{n} 2x\frac{k}{n}\binom{n}{k} x^k (1 - x)^{n-k} = -2x^2,$$

$$\sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 \binom{n}{k} x^k (1 - x)^{n-k} = x^2\left(1 - \frac{1}{n}\right) + \frac{x}{n}.$$

If we add these three equations together we have that

$$\sum_{k=0}^{n} \left(x^2 - 2x\frac{k}{n} + \left(\frac{k}{n}\right)^2\right) \binom{n}{k} x^k (1 - x)^{n-k} = x^2 - 2x^2 + x^2\left(1 - \frac{1}{n}\right) + \frac{x}{n},$$

or that

$$\sum_{k=0}^{n} \left(\frac{k}{n} - x\right)^2 \binom{n}{k} x^k (1 - x)^{n-k} = x\left(\frac{1 - x}{n}\right) \quad \text{for } 0 \leq x \leq 1.$$

Statement of Theorem

Let f be any continuous function on $[a,b]$. Then, given $\varepsilon > 0$, there is a sequence of polynomials $\{\rho_n\}_{n=1}^{\infty}$ such that $\{\rho_n\}_{n=1}^{\infty}$ converges uniformly to f on $[a,b]$.

Proof

By Chapter I we know it is sufficient to prove this theorem for the special case in which $[a,b] = [0,1]$.

For any continuous f on $[0, 1]$ we define a sequence of polynomials $\{B_n\}_{n=1}^{\infty}$ as follows:

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1 - x)^{n-k} f\left(\frac{k}{n}\right) \quad \text{for } 0 \leq x \leq 1 \text{ and } n \in I.$$

We call B_n the n^{th} Bernstein polynomial for f. Given $\varepsilon > 0$ we shall show that there exists $N \in I$ such that $|f(x) - B_n(x)| < \varepsilon$ for all $x \in [0,1]$ and $n \geq N$. This will show that $\{B_n\}_{n=1}^{\infty}$ converges uniformly to f on $[0,1]$.

From the Preliminary Computations we need the following two equations:

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1 - x)^{n-k} = 1 \tag{1}$$

and

$$\sum_{k=0}^{n} \left(\frac{k}{n} - x\right)^2 \binom{n}{k} x^k (1 - x)^{n-k} = x\left(\frac{1 - x}{n}\right) \tag{2}$$

for $0 \leq x \leq 1$ and $n \in I$.

f is uniformly continuous on $[0,1]$ since f is continuous on the closed interval $[0,1]$. Hence, given $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \frac{\varepsilon}{2}$ whenever $|x - y| < \delta$ and $x, y \in [0,1]$. Let $M = \sup|f(x)|$ for $x \in [0,1]$. We may assume $M \geq 0$. Choose $N \in I$ such that

$$\frac{1}{\sqrt[4]{N}} < \delta \tag{3}$$

and such that

$$\frac{1}{\sqrt{N}} < \frac{\varepsilon}{4M}. \tag{4}$$

Now fix $x \in [0,1]$. Multiplying (1) by $f(x)$ and subtracting B_n, we obtain for any $n \in I$

$$\begin{aligned} f(x) - B_n(x) &= \sum_{k=0}^{n} \left[f(x) - f\left(\frac{k}{n}\right)\right]\binom{n}{k}x^k(1 - x)^{n-k} \\ &= \sum{}' + \sum{}'' \end{aligned} \tag{5}$$

where $\sum'$ is the sum over those values of k such that

$$\left|\frac{k}{n} - x\right| < \frac{1}{\sqrt[4]{n}}, \tag{6}$$

while $\sum''$ is the sum over the other values of k.

If k does not satisfy (6), that is, if $\left|\frac{k}{n} - x\right| \geq \frac{1}{\sqrt[4]{n}}$, then $(k - nx)^2 = n^2\left|\frac{k}{n} - x\right|^2 \geq \sqrt{n^3}$. Hence

$$\begin{aligned} \left|\sum{}''\right| &= \left|\sum{}'' \left[f(x) - f\left(\frac{k}{n}\right)\right]\binom{n}{k}x^k(1 - x)^{n-k}\right| \\ &\leq \sum{}'' \left[|f(x)| + \left|f\left(\frac{k}{n}\right)\right|\right]\binom{n}{k}x^k(1 - x)^{n-k} \end{aligned}$$

$$\leq 2M \sum'' \binom{n}{k} x^k (1 - x)^{n-k}$$

$$\leq \frac{2M}{\sqrt{n^3}} \sum'' (k - nx)^2 \binom{n}{k} x^k (1 - x)^{n-k}$$

$$\leq \frac{2M}{\sqrt{n^3}} \sum_{k=0}^{n} (k - nx)^2 \binom{n}{k} x^k (1 - x)^{n-k}$$

$$= \frac{2M}{\sqrt{n^3}} \sum_{k=0}^{n} n^2 \left(\frac{k}{n} - x\right)^2 \binom{n}{k} x^k (1 - x)^{n-k}.$$

Multiplying (2) by n^2 we see that

$$\left|\sum''\right| \leq \frac{2M}{\sqrt{n^3}} nx(1 - x) \leq \frac{2M}{\sqrt{n}}.$$

If $n \geq N$, it follows from (4) that $\frac{1}{\sqrt{n}} < \frac{\varepsilon}{4M}$ or $\frac{2M}{\sqrt{n}} < \frac{\varepsilon}{2}$ and so

$$\left|\sum''\right| < \frac{\varepsilon}{2}.$$

Moreover, if $n \geq N$ and if k satisfies (6), then, by (3) and (6), $\left|\frac{k}{n} - x\right| < \delta$ and so

$$\left|f(x) - f\left(\frac{k}{n}\right)\right| < \frac{\varepsilon}{2}.$$

Thus

$$\begin{aligned} \left|\sum'\right| &= \left|\sum' \left[f(x) - f\left(\frac{k}{n}\right)\right] \binom{n}{k} x^k (1 - x)^{n-k}\right| \\ &< \frac{\varepsilon}{2} \sum' \binom{n}{k} x^k (1 - x)^{n-k} \end{aligned}$$

and so by (1)

$$\left|\sum'\right| < \frac{\varepsilon}{2}.$$

Thus, from (5),

$$|f(x) - B_n(x)| = |\sum' + \sum''| \le |\sum'| + |\sum''| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Since x was any point in $[0,1]$ and n any integer with $n \ge N$, this shows $|f(x) - B_n(x)| < \varepsilon$ for $0 \le x \le 1$ and $n \ge N$.

Remarks

The idea for this proof arises from probability, as follows: Let S_n be the number of successes in n independent trials with the probability of success in each being p $(0 < p < 1)$. Then the weak law of large numbers says that

$$\mathrm{Prob}\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right) \to 0 \quad \text{as} \quad n \to \infty \quad \text{for each fixed} \quad \varepsilon,$$

or $\frac{S_n}{n} \to p$ in probability.

Now if f is a real arbitrary function, it seems reasonable that $f\left(\frac{S_n}{n}\right) \to f(p)$ in probability, and furthermore that $E\left(f\left(\frac{S_n}{n}\right)\right) \to E(f(p)) = f(p)$. But

$$E\left(f\left(\frac{S_n}{n}\right)\right) = \sum_{k=1}^{n} f\left(\frac{k}{n}\right)\mathrm{Prob}(S_n = k)$$

$$= \sum_{k=1}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}p^k(1 - p)^{n-k}.$$

With this in mind, Bernstein justified these steps, and then showed the convergence is uniform.

V. THE STONE-WEIERSTRASS THEOREM

In this final chapter we present a generalization of the Weierstrass Approximation Theorem called the Stone-Weierstrass Theorem. This theorem is helpful to those who wish to study operator theory and functional analysis.

We will prove this theorem using the following three steps. First we will show that A, a closed subalgebra of $C(X,R)$, for X a compact Hausdorff space, is also a closed sublattice. Next we will show that if A separates points and contains a nonzero constant function, it strongly separates points. Finally we show that A equals $C(X,R)$.

The Stone-Weierstrass Theorem

Let X be a compact Hausdorff space and let A be a closed subalgebra of $C(X,R)$ which separates points and contains a nonzero constant function. Then A equals $C(X,R)$.

Proof

Note that $C(X,R)$, the set of all real continuous functions on X, is a Banach algebra.

STEP I: Let X be a compact Hausdorff space. Let A be a closed subalgebra of $C(X,R)$. Then A is a closed sublattice of $C(X,R)$.

In the proof of this step we make use of the concept of the absolute value of a function. If f is a real

function defined on a topological space X, then the function $|f|$ is defined by $|f|(x) = |f(x)|$. If f is continuous, then $|f|$ is also continuous. We observe that the lattice operations in $C(X,R)$ are expressible in terms of addition, scalar multiplication, and the formation of absolute values:

$$f \vee g = \frac{f + g + |f - g|}{2} = \max\{f,g\} \quad \text{and} \quad f \wedge g = \frac{f + g - |f - g|}{2} = \min\{f,g\}.$$

These identities show that any linear subspace (that is, a subspace closed under addition and scalar multiplication) of $C(X,R)$ which contains the absolute value of each of its functions is a sublattice of $C(X,R)$.

By the above remarks it suffices to show that if f is in A, then $|f|$ is also in A. Let $\varepsilon > 0$ be given. Since $|t|$ is a continuous function of the real variable t, by the Weierstrass Approximation Theorem there exists a polynomial p' with the property that $\left||t| - p'(t)\right| < \frac{\varepsilon}{2}$ for every t on the closed interval $[-\|f\|,\|f\|]$. Let p be the polynomial which results when the constant term of p' is replaced by zero. Then p is a polynomial with zero as its constant term. Since $|0| = 0$, and $\left||t| - p'(t)\right| < \frac{\varepsilon}{2}$ or $|0 - p'(0)| = |p'(0)| < \frac{\varepsilon}{2}$, the constant term of p' must be less than $\frac{\varepsilon}{2}$. Then, when we replace the constant term in p' by zero, we are changing p' at most by $\frac{\varepsilon}{2}$, so that the largest difference we could now have between p and $|t|$ is $\frac{\varepsilon}{2} + \frac{\varepsilon}{2}$ or ε. Thus $\left||t| - p(t)\right| < \varepsilon$ for every t in $[-\|f\|,\|f\|]$. Since A is an algebra, the

function $p(f)$ in $C(X,R)$ is in A. $f(x)$ for any fixed x is just a real number, in fact, a real number in $[-\|f\|,\|f\|]$. Thus it is easy to see that $\left||f(x)| - p(f(x))\right| < \varepsilon$ for every point x in X, and from this it follows that $\||f| - p(f)\| < \varepsilon$. We conclude the proof by remarking that since A is closed, the fact that $|f|$ can be approximated by the function $p(f)$ in A shows that $|f|$ is a limit point of A and thus is in A.

STEP II: Show that A strongly separates points; that is, if x and y are any two distinct points of X, and if a and b are any two real numbers, then there exists $f \in A$ such that $f(x) = a$ and $f(y) = b$.

Note that if X has only one point, then $C(X,R)$ contains only constant functions, and since A contains a nonzero constant function and is a subalgebra, it contains all constant functions, and thus equals $C(X,R)$. We may thus assume that X has more than one point.

Since A separates points, there exists a function g in A such that $g(x) \neq g(y)$, where x and y are distinct points of X. We now define f by

$$f(z) = a\frac{g(z) - g(y)}{g(x) - g(y)} + b\frac{g(z) - g(x)}{g(y) - g(x)}.$$

Since $g(x) - g(y)$ is just a nonzero real number, f is in A, and

$$f(x) = a\frac{g(x) - g(y)}{g(x) - g(y)} + b\frac{g(x) - g(x)}{g(y) - g(x)} = a, \text{ and}$$

$$f(y) = a\frac{g(y) - g(y)}{g(x) - g(y)} + b\frac{g(y) - g(x)}{g(y) - g(x)} = b.$$

Thus f has the required properties.

STEP III: Show $A = C(X,R)$.

Let f be an arbitrary function in $C(X,R)$. Since $A \subset C(X,R)$, we need only show $C(X,R) \subset A$, that is, that f is in A, to show that they are equal.

Since A is closed, it contains all of its limit points. We will show that f is a limit point of A; that is, for $\varepsilon > 0$, there exists a function g in A such that $\|f - g\| < \varepsilon$. Thus what we really have to prove is that for $\varepsilon > 0$ there exists g in A such that $f(z) - \varepsilon < g(z) < f(z) + \varepsilon$ for all z in X. We now construct such a function.

Let x be a point in X which is fixed, and let y be a point in X different from x. By Step II, there exists a function f_y in A such that $f_y(x) = f(x)$ and $f_y(y) = f(y)$. Now consider the open set $G_y = \{z \in X: f_y(z) < f(z) + \varepsilon\}$. Both x and y belong to G_y, for $f_y(x) = f(x)$ implies that $f_y(x) < f(x) + \varepsilon$, and similarly for y, so the class of G_y's for all points y different from x is an open cover of X. Since X is compact (by hypothesis), this open cover has a finite subcover, which we denote by $\{G_{y_1}, G_{y_2}, \ldots, G_{y_n}\}$. If the corresponding functions in A are denoted by

$f_{y_1}, f_{y_2}, \ldots, f_{y_n}$, then $g_x = f_{y_1} \wedge f_{y_2} \wedge \ldots \wedge f_{y_n}$ is a function in A such that $g_x(x) = f(x)$. We see that if $z \in G_{y_k}$, then

$g_x(z) = \min\{f_{y_1}(z), \ldots, f_{y_n}(z)\} \leq f_{y_k}(z) < f(z) + \varepsilon$. Thus $g_x(z) \leq f(z) + \varepsilon$ for all points z in X.

We next consider the open set $H_x = \{z \in X: g_x(z) > f(z) - \varepsilon\}$. Since x belongs to H_x, the class of H_x's for all points x in X is an open cover of X. The compactness of X implies that this open cover has a finite subcover, which we denote by $\{H_{x_1}, H_{x_2}, \ldots, H_{x_m}\}$. We denote the corresponding functions in A by $g_{x_1}, g_{x_2}, \ldots, g_{x_m}$, and we define g by $g = g_{x_1} \vee g_{x_2} \vee \ldots \vee g_{x_m}$. We see that if $z \in H_{x_k}$, then $g(z) = \max\{g_{x_1}(z), \ldots, g_{x_m}(z)\} \geq g_{x_k}(z) > f(z) - \varepsilon$.

Thus it is clear that g is a function in A with the property that $f(z) - \varepsilon < g(z) < f(z) + \varepsilon$ for all points z in X, so our proof is complete.

<u>Comment</u>

This theorem does not hold in general for complex algebras. A counterexample would be:

> Let A be the set of all continuous functions in $|z| \leq 1$ which are analytic in $|z| < 1$. We see that A is an algebra, and since the function

$g(z) = z$ is in A, A separates points. However, A is not $C(|z| \leq 1, C)$, since $f(z) = |z|$ is not analytic.

However, the conclusion of the theorem does hold, even for complex algebras, if an extra condition is imposed on A; namely, that for every $f \in A$, its complex conjugate $\bar{f}$ must also belong to A.

BIBLIOGRAPHY

GOLDBERG, R. R.: "Methods of Real Analysis," Xerox College Publishing, Lexington, Mass., 1964.

RICE, J. R.: "The Approximation of Functions," Addison-Wesley Publishing Company, Reading, Mass., 1964.

RUDIN, W.: "Principles of Mathematical Analysis," McGraw-Hill Book Company, New York, 1964.

SIMMONS, G. F.: "Introduction to Topology and Modern Analysis," McGraw-Hill Book Company, New York, 1963.

DATE
ILMED
8